

Lecture 3: Linear Algebra

Linear algebra studies *linear systems*: systems of equations in which every equation is linear. It is the foundation of optimization and also a powerful tool for analyzing nonlinear systems. In this lecture we introduce the basic vocabulary and operations of linear algebra, and we study the solution sets of linear systems. Throughout this lecture, vectors are columns.

1 Matrix algebra

1.1 Matrix vocabulary and basic arithmetic

Definition 1.1 (Matrices and transpose). An $m \times n$ *matrix* $A = (a_{ij})$ is a rectangular array with m rows and n columns, with a_{ij} in row i and column j . Two matrices are equal when they have the same dimensions and all corresponding entries are equal. The *transpose* of A is the $n \times m$ matrix $A^\top$ defined by

$$(A^\top)_{ij} = a_{ji}.$$

A matrix is *square* if it has the same number of rows and columns, *diagonal* if every off-diagonal entry is zero, and *symmetric* if $A = A^\top$.

Matrices of the same dimensions are added entrywise, and scalar multiplication is entrywise:

$$(A + B)_{ij} = a_{ij} + b_{ij}, \quad (\alpha A)_{ij} = \alpha a_{ij}.$$

The $m \times n$ *zero matrix* has every entry zero. The $n \times n$ *identity matrix* I_n has ones on the diagonal and zeros elsewhere. The diagonal matrix with diagonal entries $d_1, \dots, d_n$ is written $\text{diag}(d_1, \dots, d_n)$.

Example 1.2 (Matrix arithmetic). Let

$$A = \begin{pmatrix} 1 & 2 & 0 \\ -1 & 0 & 3 \end{pmatrix}, \quad B = \begin{pmatrix} 0 & 1 & 1 \\ 2 & -1 & 0 \end{pmatrix}.$$

Then

$$A + B = \begin{pmatrix} 1 & 3 & 1 \\ 1 & -1 & 3 \end{pmatrix}, \quad 2A = \begin{pmatrix} 2 & 4 & 0 \\ -2 & 0 & 6 \end{pmatrix}, \quad A^\top = \begin{pmatrix} 1 & -1 \\ 2 & 0 \\ 0 & 3 \end{pmatrix}.$$

Addition requires matching dimensions. Multiplication uses a different compatibility rule.

1.2 Matrix multiplication

Definition 1.3 (Matrix multiplication). If A is $m \times n$ and B is $n \times p$, then their product AB is the $m \times p$ matrix with

$$(AB)_{ij} = \sum_{k=1}^n a_{ik}b_{kj}.$$

The product is defined only when the inner dimensions agree. In that case the matrices are *conformable*:

$$(m \times n)(n \times p) = (m \times p).$$

Each entry of AB is a row of A multiplied against a column of B .

Example 1.4 (Matrix multiplication). Let

$$C = \begin{pmatrix} 1 & 2 & 0 \\ 0 & 1 & 1 \end{pmatrix}, \quad D = \begin{pmatrix} 1 & 0 \\ 0 & 1 \\ 1 & 1 \end{pmatrix}.$$

Then C is 2×3 and D is 3×2 , so

$$CD = \begin{pmatrix} 1 & 2 \\ 1 & 2 \end{pmatrix}$$

is 2×2 . The reverse product is also defined in this example, but it is different in both size and entries:

$$DC = \begin{pmatrix} 1 & 2 & 0 \\ 0 & 1 & 1 \\ 1 & 3 & 1 \end{pmatrix}.$$

Since $CD \neq DC$, this example also shows that matrix multiplication is not generally commutative.

Whenever the products are conformable,

$$(AB)C = A(BC), \quad A(B + C) = AB + AC, \quad (AB)^\top = B^\top A^\top,$$

and the identity matrices satisfy $AI_n = A$ and $I_m A = A$ for an $m \times n$ matrix A .

Vectors can be treated as matrices with one column. For $x, y \in \mathbb{R}^n$, $x^\top y$ is a 1×1 matrix whose entry is the inner product $x \cdot y$, whereas $xy^\top$ is an $n \times n$ matrix.

1.3 Linear combinations, span, and independence

Definition 1.5 (Linear combination). Let $v_1, \dots, v_k \in \mathbb{R}^n$ be a collection of vectors. A *linear combination* of $v_1, \dots, v_k$ is a vector $\sum_{i=1}^k \alpha_i v_i$ with $\alpha_i \in \mathbb{R}$.

Below is a collection of vectors we will use to illustrate concepts in this lecture.

Example 1.6 (A linear combination). Let $v_1 = \begin{pmatrix} 1 \\ 2 \end{pmatrix}$, $v_2 = \begin{pmatrix} 2 \\ 2 \end{pmatrix}$, and $v_3 = \begin{pmatrix} 3 \\ 4 \end{pmatrix}$. Here $v_3 = v_1 + v_2$, so v_3 is a linear combination of v_1 and v_2 , with $\alpha_1 = \alpha_2 = 1$.

Definition 1.7 (Subspace and span). Let $V \subseteq \mathbb{R}^n$ and let $v_1, \dots, v_k \in \mathbb{R}^n$.

- The set V is a *subspace* of $\mathbb{R}^n$ if it is nonempty and

$$x, y \in V, \quad \alpha, \beta \in \mathbb{R} \quad \implies \quad \alpha x + \beta y \in V.$$

- Their *span* is the set of all such combinations:

$$\text{span}\{v_1, \dots, v_k\} = \left\{ \sum_{i=1}^k \alpha_i v_i : \alpha_1, \dots, \alpha_k \in \mathbb{R} \right\}.$$

Every span is a subspace. Indeed, adding two linear combinations of $v_1, \dots, v_k$, or multiplying one by a scalar, produces another linear combination of the same vectors.

The span of one nonzero vector v is the line $\{\alpha v : \alpha \in \mathbb{R}\}$ through the origin. A line that does not go through the origin cannot be a subspace, because every subspace contains the zero vector.

Definition 1.8 (Independence, basis, and dimension). Let $V \subseteq \mathbb{R}^n$ be a subspace.

- The collection $v_1, \dots, v_k$ of vectors is *linearly independent* if

$$\alpha_1 v_1 + \dots + \alpha_k v_k = 0 \quad \implies \quad \alpha_1 = \dots = \alpha_k = 0$$

for all $\alpha_1, \dots, \alpha_k \in \mathbb{R}$. Otherwise it is *linearly dependent*.

- A *basis* of V is a linearly independent collection of vectors that spans V .
- The *dimension* $\dim V$ of V is the number of vectors in a basis of V .

In $\mathbb{R}^3$ we can illustrate all four possible dimensions of subspaces — 0, 1, 2, and 3. That is, every subspace of $\mathbb{R}^3$ is one of the following: the origin; a line through the origin; a plane through the origin; or all of $\mathbb{R}^3$. Figure 1.1 shows one subspace of each dimension. This is the picture to keep in mind for any subspace of $\mathbb{R}^n$: a subspace is flat, passes through the origin, and has dimension at most n . Unless it is $\{0\}$, it extends without bound in every direction that it contains.

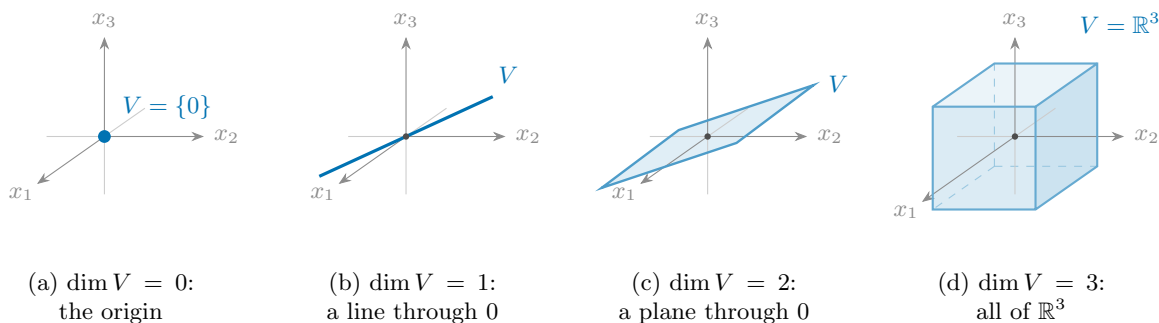


Figure 1.1: The four kinds of subspaces of $\mathbb{R}^3$, one for each dimension: the origin alone, a line through the origin, a plane through the origin, and the whole space. Every subspace of $\mathbb{R}^3$ is one of these. Each is flat and contains the origin. The line and plane panels show representative viewing windows: those subspaces continue without bound. There is one dimension-0 subspace and one dimension-3 subspace, but infinitely many dimension-1 and dimension-2 subspaces.

One could think of linear dependence as redundancy: if vectors are dependent, then at least one of them can be expressed as a linear combination of the others.

Example 1.9. Let

$$v_1 = \begin{pmatrix} 1 \\ 2 \end{pmatrix}, \quad v_2 = \begin{pmatrix} 2 \\ 2 \end{pmatrix}, \quad v_3 = \begin{pmatrix} 3 \\ 4 \end{pmatrix}.$$

Because $v_3 = v_1 + v_2$, the three-vector collection is dependent and v_3 may be deleted without changing its span. The remaining vectors v_1 and v_2 are independent: if $\alpha_1 v_1 + \alpha_2 v_2 = 0$, then

$$\alpha_1 + 2\alpha_2 = 0, \quad 2\alpha_1 + 2\alpha_2 = 0,$$

which forces $\alpha_1 = \alpha_2 = 0$. Thus v_1, v_2 form a basis of $\mathbb{R}^2$, and the three vectors span $\mathbb{R}^2$. Figure 1.2 illustrates these relationships.

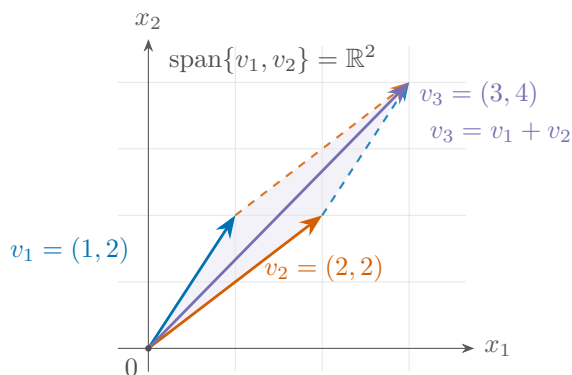


Figure 1.2: The vectors v_1, v_2, v_3 in $\mathbb{R}^2$. The solid arrows all begin at the origin. The dashed translated copies complete the shaded parallelogram and show $v_3 = v_1 + v_2$. Thus v_3 contributes no new direction: v_1, v_2 are independent and already span $\mathbb{R}^2$, while the three-vector collection is dependent. The span is the entire plane, not only the shaded parallelogram.

We use the standard finite-dimensional facts that every subspace of $\mathbb{R}^n$ has a basis, any two bases of the same subspace have the same number of vectors, and every independent collection in $\mathbb{R}^n$ has at most n vectors. The *standard basis* of $\mathbb{R}^n$ is the collection $e_1, \dots, e_n$, where e_i is the vector with 1 in coordinate i and 0 in every other coordinate.

2 Linear maps and linear systems

So far we have treated matrices as arrays of numbers. We now show that matrices represent linear functions from $\mathbb{R}^n$ to $\mathbb{R}^m$, and we use this perspective to study the solution sets of systems of linear equations. The same product Ax also has two useful interpretations. We begin by naming the rows and columns of A , and then develop each interpretation.

2.1 Two readings of a matrix-vector product

Let

$$A = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix}.$$

Write the rows of A as $r_1, r_2, \dots, r_m$ and its columns as $c_1, c_2, \dots, c_n$. Thus

$$A = \begin{pmatrix} r_1 \\ r_2 \\ \vdots \\ r_m \end{pmatrix} = \begin{pmatrix} c_1 & c_2 & \cdots & c_n \end{pmatrix}.$$

Suppose $x \in \mathbb{R}^n$. We call x the *input* and $Ax \in \mathbb{R}^m$ the *output*. The product can be represented as

$$Ax = \begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & & \vdots \\ a_{m1} & \cdots & a_{mn} \end{pmatrix} \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} a_{11}x_1 + \cdots + a_{1n}x_n \\ \vdots \\ a_{m1}x_1 + \cdots + a_{mn}x_n \end{pmatrix} = x_1c_1 + \cdots + x_nc_n. \quad (1)$$

The product Ax is thus a linear combination of the columns of A , with entries of x as weights. For the row interpretation of matrix A , we need to consider an equation $Ax = b$.

Definition 2.1 (Linear system). For an $m \times n$ matrix A and $b \in \mathbb{R}^m$, the equation

$$Ax = b, \quad x \in \mathbb{R}^n,$$

is a *linear system* with *coefficient matrix* A and *right-hand side* b . It is *consistent* if it has a solution and *inconsistent* otherwise.

The row interpretation writes the system as one scalar equation for each row of A :

$$Ax = b \iff r_ix = a_{i1}x_1 + \cdots + a_{in}x_n = b_i, \quad \text{for each row } i. \quad (2)$$

Each row is thus a restriction on what the input x could be to reach output b using matrix A . At the same time, the column representation essentially asks which linear combination of the columns of A produces the target b .

Make sure you understand the representations in (1) and (2), as they will be crucial for understanding just about everything that comes next. Here is the running example.

Example 2.2. Let

$$A = \begin{pmatrix} 1 & 2 & 3 \\ 2 & 2 & 4 \end{pmatrix}, \quad x = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}.$$

The rows and columns of A are

$$r_1 = (1 \ 2 \ 3), \quad r_2 = (2 \ 2 \ 4),$$

and

$$c_1 = \begin{pmatrix} 1 \\ 2 \end{pmatrix}, \quad c_2 = \begin{pmatrix} 2 \\ 2 \end{pmatrix}, \quad c_3 = \begin{pmatrix} 3 \\ 4 \end{pmatrix}.$$

The column interpretation gives

$$Ax = x_1c_1 + x_2c_2 + x_3c_3 = x_1 \begin{pmatrix} 1 \\ 2 \end{pmatrix} + x_2 \begin{pmatrix} 2 \\ 2 \end{pmatrix} + x_3 \begin{pmatrix} 3 \\ 4 \end{pmatrix}.$$

For

$$b = \begin{pmatrix} 3 \\ 4 \end{pmatrix},$$

the row interpretation writes $Ax = b$ as

$$x_1 + 2x_2 + 3x_3 = 3, \quad 2x_1 + 2x_2 + 4x_3 = 4.$$

The theorem below shows that a function from $\mathbb{R}^n$ to $\mathbb{R}^m$ is linear if and only if it can be represented as an $m \times n$ matrix. We use the word map or transformation instead of the word function to emphasize that the domain and codomain are vector spaces, and that the function preserves the vector space structure.

Definition 2.3 (Linear map). A function $T : \mathbb{R}^n \rightarrow \mathbb{R}^m$ is a *linear map*, or a *linear transformation*, if

$$T(x + y) = T(x) + T(y), \quad T(\alpha x) = \alpha T(x)$$

for all $x, y \in \mathbb{R}^n$ and $\alpha \in \mathbb{R}$.

Here x is the *input* and $T(x)$ is the *output*.

Theorem 2.4 (Matrices represent linear maps). *A function $T : \mathbb{R}^n \rightarrow \mathbb{R}^m$ is linear if and only if there is an $m \times n$ matrix A such that $T(x) = Ax$ for every $x \in \mathbb{R}^n$. In that case A is unique, and its j th column is $T(e_j)$. If $S : \mathbb{R}^m \rightarrow \mathbb{R}^p$ is linear with matrix B , then the composition $S \circ T$ is linear with matrix BA .*

Proof (optional). Linear implies matrix. Let T be linear. Every $x \in \mathbb{R}^n$ is the linear combination $x = x_1 e_1 + \cdots + x_n e_n$ of the standard basis, so linearity applied to each term gives

$$T(x) = T\left(\sum_{j=1}^n x_j e_j\right) = \sum_{j=1}^n x_j T(e_j).$$

Let $A = [T(e_1) \cdots T(e_n)]$, which is $m \times n$ because each $T(e_j)$ lies in $\mathbb{R}^m$. Since Ax is the combination of the columns of A with weights $x_1, \dots, x_n$, the display says exactly that $T(x) = Ax$.

Matrix implies linear. If $T(x) = Ax$ for an $m \times n$ matrix A , then $A(x + y) = Ax + Ay$ and $A(\alpha x) = \alpha(Ax)$, so T is linear.

Uniqueness. Suppose $Ax = Bx$ for every $x \in \mathbb{R}^n$. Taking $x = e_j$ makes Ae_j and Be_j the j th columns of A and B , so the two matrices agree column by column and $A = B$. The same substitution in $T(x) = Ax$ identifies the j th column of A as $T(e_j)$.

Composition. Let S be linear with matrix B . For every $x \in \mathbb{R}^n$, associativity of the matrix product gives

$$(S \circ T)(x) = S(T(x)) = B(Ax) = (BA)x,$$

so $S \circ T$ is represented by the $p \times n$ matrix BA , and it is linear by the second part. $\square$

2.2 Row space, column space, and rank

Definition 2.5 (Row and column spaces). Let A be an $m \times n$ matrix. Its *row space* and *column space* are

$$\text{Row}(A) = \text{span}\{r_1, \dots, r_m\} \subseteq \mathbb{R}^n, \quad \text{Col}(A) = \text{span}\{c_1, \dots, c_n\} \subseteq \mathbb{R}^m.$$

By (1),

$$\text{Col}(A) = \{Ax : x \in \mathbb{R}^n\},$$

so the column space is the range of the linear map $x \mapsto Ax$.

The row and column spaces live in different ambient spaces, but they have the same dimension. We use this fundamental result without proof.

Theorem 2.6 (Equality of row and column rank). *For every matrix A ,*

$$\dim \text{Row}(A) = \dim \text{Col}(A).$$

Definition 2.7 (Rank). The *rank* of A is their common dimension:

$$\text{rank } A = \dim \text{Row}(A) = \dim \text{Col}(A).$$

The row reading interprets rank as the number of independent restrictions. The column reading interprets it as the dimension of the set of outputs that A can produce.

Example 2.8 (Rank of a matrix). Let

$$A = \begin{pmatrix} 1 & 2 & 3 \\ 2 & 2 & 4 \end{pmatrix}.$$

The first two columns of A form a basis of $\mathbb{R}^2$. Thus

$$\text{Col}(A) = \mathbb{R}^2, \quad \text{rank } A = 2.$$

The two rows are consequently independent: the system contains two independent restrictions on three variables.

2.3 Null space and changes that preserve the restrictions

Suppose the input x produces the output b , so $Ax = b$. Let $h \in \mathbb{R}^n$ be a proposed change to the input. After the change, the new input is $x + h$. It produces the same output exactly when

$$A(x + h) = b \iff Ah = 0.$$

Thus h leaves Ax unchanged if and only if $Ah = 0$. The collection of all such changes is the null space.

Definition 2.9 (Null space). For an $m \times n$ matrix A , its *null space*, or *kernel*, is

$$N(A) = \{h \in \mathbb{R}^n : Ah = 0\}.$$

The dimension of $N(A)$ is called the *nullity* of A .

Every null space is a subspace. Indeed, if $u, v \in N(A)$ and $\alpha, \beta \in \mathbb{R}$, then

$$A(\alpha u + \beta v) = \alpha Au + \beta Av = 0.$$

We have now seen two ways to specify a subspace: a span specifies one using generating vectors, whereas a null space specifies one using equations of the form $Ax = 0$.

Its elements have two interpretations. In the row interpretation,

$$Ah = 0 \iff a_{i1}h_1 + \cdots + a_{in}h_n = 0 \quad \text{for every row } i,$$

so adding h to x changes the left-hand side of every constraint by zero. In the column reading,

$$Ah = 0 \iff h_1c_1 + \cdots + h_nc_n = 0.$$

Here h_j is the change in the weight on column c_j . Thus $Ah = 0$ means that these changes cancel, leaving the output unchanged. A nonzero such h exists exactly when the columns are linearly dependent.

Rank and nullity measure two different things. The rank is the dimension of the set of outputs that A can produce. The nullity is the dimension of the set of changes that leave Ax unchanged. The rank–nullity theorem relates these two counts. We use it without proof.

Theorem 2.10 (Rank–nullity theorem). *If A is an $m \times n$ matrix, then*

$$\text{rank } A + \dim N(A) = n.$$

Example 2.11 (Null space of a matrix). Let

$$A = \begin{pmatrix} 1 & 2 & 3 \\ 2 & 2 & 4 \end{pmatrix}.$$

Here $n = 3$ and $\text{rank } A = 2$, so $\dim N(A) = 1$. To identify the null space, solve $Ah = 0$:

$$h_1 + 2h_2 + 3h_3 = 0, \quad 2h_1 + 2h_2 + 4h_3 = 0.$$

Subtracting twice the first equation from the second gives $h_2 = -h_3$. Substitution into the first gives $h_1 = -h_3$. Letting $t = -h_3$,

$$N(A) = \left\{ \begin{pmatrix} t \\ t \\ -t \end{pmatrix} : t \in \mathbb{R} \right\} = \text{span} \left\{ \begin{pmatrix} 1 \\ 1 \\ -1 \end{pmatrix} \right\}.$$

This illustrates rank–nullity:

$$\text{rank } A + \dim N(A) = 2 + 1 = 3 = n.$$

Here the nullity 1 means that one scalar, t , can be chosen freely when describing changes that preserve all restrictions. More generally, if $\text{rank } A = r$, then $n - r$ coefficients can be chosen freely; these are the $n - r$ degrees of freedom.

2.4 The solution set

The column interpretation answers the existence question immediately:

$$Ax = b \text{ is consistent} \iff b \in \text{Col}(A). \quad (3)$$

The following theorem describes the solution set of a consistent linear system $Ax = b$ in terms of a particular solution and the null space (which is the solution set for $Ax = 0$).

Theorem 2.12 (Solution set). *If x_0 is one solution of $Ax = b$, then the full solution set is*

$$\{x \in \mathbb{R}^n : Ax = b\} = x_0 + N(A), \quad \text{where } x_0 + N(A) = \{x_0 + h : h \in N(A)\}.$$

Consequently, a consistent system has a unique solution if and only if $N(A) = \{0\}$.

Proof. If $Ax = b = Ax_0$, then $A(x - x_0) = 0$, so $x - x_0 \in N(A)$. Conversely, if $h \in N(A)$, then $A(x_0 + h) = Ax_0 + Ah = b$. $\square$

Together with the column-space criterion, the theorem gives the complete description

$$\{x \in \mathbb{R}^n : Ax = b\} = \begin{cases} \emptyset, & b \notin \text{Col}(A), \\ x_0 + N(A), & b \in \text{Col}(A). \end{cases}$$

Example 2.13 (Solution set of a linear system). Let

$$A = \begin{pmatrix} 1 & 2 & 3 \\ 2 & 2 & 4 \end{pmatrix}, \quad b = \begin{pmatrix} 3 \\ 4 \end{pmatrix}.$$

The null space is

$$N(A) = \text{span} \left\{ \begin{pmatrix} 1 \\ 1 \\ -1 \end{pmatrix} \right\},$$

and one particular solution is $x_0 = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}$. Therefore

$$\{x : Ax = b\} = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} + \text{span} \left\{ \begin{pmatrix} 1 \\ 1 \\ -1 \end{pmatrix} \right\}.$$

Adding the displayed null vector gives another solution:

$$\begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} + \begin{pmatrix} 1 \\ 1 \\ -1 \end{pmatrix} = \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix},$$

reflecting that the third column of A is the sum of its first two columns.

2.5 Rank, existence, and uniqueness

The preceding results give the two full-rank criteria.

Theorem 2.14 (Full row and column rank). *Let A be $m \times n$.*

1. *rank $A = m$ if and only if $Ax = b$ has at least one solution for every $b \in \mathbb{R}^m$.*
2. *rank $A = n$ if and only if $Ax = b$ has at most one solution for every $b \in \mathbb{R}^m$.*

The first clause follows because rank $A = m$ exactly when $\text{Col}(A) = \mathbb{R}^m$, which means every $b \in \mathbb{R}^m$ belongs to the column space. For the second clause, rank-nullity gives

$$\text{rank } A = n \iff N(A) = \{0\}.$$

The solution-set theorem then says that a consistent system has exactly one solution.

Together, rank and the column-space criterion classify every system:

Rank condition	Solutions
rank $A = m = n$	every b has exactly one
rank $A = m < n$	every b has infinitely many
rank $A = n < m$	none or exactly one, depending on b
rank $A < \min\{m, n\}$	none or infinitely many, depending on b

3 Square matrices and invertibility

Definition 3.1 (Inverse). An $n \times n$ matrix A is *invertible*, or *nonsingular*, if there is an $n \times n$ matrix A^{-1} satisfying

$$A^{-1}A = AA^{-1} = I_n.$$

If no inverse exists, A is *singular*.

An inverse is unique: if B and C are both inverses of A , then $B = B(AC) = (BA)C = C$.

Theorem 3.2 (Invertible-matrix equivalences). *For an $n \times n$ matrix A , the following are equivalent:*

1. *A is invertible;*
2. *rank $A = n$;*
3. *the columns of A are linearly independent;*
4. *$N(A) = \{0\}$;*
5. *for every $b \in \mathbb{R}^n$, the system $Ax = b$ has exactly one solution.*

When these conditions hold, the solution is $x = A^{-1}b$.

The equivalence of clauses (2)–(5) follows from full row and column rank and the affine solution-set theorem. If A is invertible, multiplying $Ax = b$ by A^{-1} gives the unique solution $x = A^{-1}b$. Conversely, if every $Ax = e_j$ has a unique solution x^j , put those solutions into the columns of $B = [x^1 \cdots x^n]$. Then $AB = I_n$, and the triviality of $N(A)$ forces $BA = I_n$, so $B = A^{-1}$.

The determinant is a function that turns a square matrix into a scalar.

Definition 3.3 (Determinant). For a 1×1 matrix, $\det(a) = a$. For $n \geq 2$, let A_{1j} denote the $(n-1) \times (n-1)$ matrix obtained from A by deleting row 1 and column j . The *determinant* of A is

$$\det A = \sum_{j=1}^n (-1)^{1+j} a_{1j} \det A_{1j}.$$

This is called *cofactor expansion along the first row*: each entry of the first row is multiplied by the determinant of the matrix left after deleting that entry's row and column, and the signs alternate. Expanding along any other row or column gives the same number.

Example 3.4 (2×2 determinant). Let $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$. Here $A_{11} = (d)$ and $A_{12} = (c)$, so

$$\det A = ad - bc.$$

Example 3.5 (3×3 determinant). Let $A = \begin{pmatrix} a & b & c \\ d & e & f \\ g & h & i \end{pmatrix}$. Each of the three terms now needs a 2×2 determinant:

$$\det A = a(ei - fh) - b(di - fg) + c(dh - eg).$$

The determinant supplies a scalar test for the equivalent conditions above.

Theorem 3.6 (Determinant criterion). *A square matrix A is invertible if and only if $\det A \neq 0$.*

Example 3.7 (The inverse in the 2×2 case). Let

$$A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}, \quad \det A = ad - bc.$$

If $\det A \neq 0$, multiplying in either order gives

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix} \begin{pmatrix} d & -b \\ -c & a \end{pmatrix} = (ad - bc)I_2.$$

As long as the determinant $ad - bc \neq 0$, we can divide by it to get the inverse:

$$A^{-1} = \frac{1}{ad - bc} \begin{pmatrix} d & -b \\ -c & a \end{pmatrix}.$$

4 Eigenvalues and eigenvectors

Definition 4.1 (Eigenvalues and eigenvectors). Let A be a real $n \times n$ matrix. A scalar $\lambda \in \mathbb{R}$ is an *eigenvalue* of A if there is a nonzero vector $v \in \mathbb{R}^n$ such that

$$Av = \lambda v.$$

Such a vector v is an *eigenvector* associated with λ .

The eigenvalue equation can be rewritten as

$$Av = \lambda v \iff (A - \lambda I)v = 0. \quad (4)$$

As such, we immediately have the following equivalences, which are useful for computing eigenvalues and eigenvectors.

Theorem 4.2 (Eigenvalue equivalences). *For a real $n \times n$ matrix A and $\lambda \in \mathbb{R}$, the following are equivalent:*

1. λ is an eigenvalue of A ;
2. $N(A - \lambda I) \neq \{0\}$;
3. $A - \lambda I$ is singular;
4. $\det(A - \lambda I) = 0$.

Proof (optional). (1) $\Leftrightarrow$ (2). By Definition 4.1, λ is an eigenvalue exactly when $Av = \lambda v$ for some $v \neq 0$. By (4) such a v satisfies $(A - \lambda I)v = 0$, and conversely. So an eigenvalue exists precisely when $A - \lambda I$ has a nonzero null vector, which is the statement $N(A - \lambda I) \neq \{0\}$.

(2) $\Leftrightarrow$ (3). Apply Theorem 3.2 to the square matrix $A - \lambda I$: it is invertible if and only if its null space is trivial. Negating both sides, $A - \lambda I$ is singular if and only if $N(A - \lambda I) \neq \{0\}$.

(3) $\Leftrightarrow$ (4). Apply Theorem 3.6 to $A - \lambda I$: it is invertible if and only if $\det(A - \lambda I) \neq 0$. Negating both sides gives the claim. $\square$

If one of the eigenvalues of a matrix is zero, then the matrix is singular; that can be seen by taking $\lambda = 0$ in Theorem 4.2.

Corollary 4.3 (Zero eigenvalue). *A square matrix A is singular if and only if 0 is an eigenvalue of A . Equivalently, A is invertible if and only if all of its eigenvalues are nonzero.*

The equation

$$\det(A - \lambda I) = 0$$

is the *characteristic equation*. Its left-hand side is the *characteristic polynomial*. The roots of the characteristic polynomial are the eigenvalues of A . For each eigenvalue λ , the associated eigenvectors are the nonzero elements of $N(A - \lambda I)$.

Example 4.4 (Eigenvalues and eigenvectors of a 2×2 matrix). Let

$$A = \begin{pmatrix} 1 & 2 \\ 4 & 3 \end{pmatrix}.$$

Subtracting λ from each diagonal entry and taking the determinant gives the characteristic polynomial

$$\det(A - \lambda I) = \det \begin{pmatrix} 1 - \lambda & 2 \\ 4 & 3 - \lambda \end{pmatrix} = (1 - \lambda)(3 - \lambda) - 8 = \lambda^2 - 4\lambda - 5 = (\lambda - 5)(\lambda + 1).$$

Its roots are the eigenvalues $\lambda_1 = 5$ and $\lambda_2 = -1$.

To find the eigenvectors belonging to $\lambda_1 = 5$, solve $(A - 5I)v = 0$:

$$A - 5I = \begin{pmatrix} -4 & 2 \\ 4 & -2 \end{pmatrix},$$

so both equations reduce to $v_2 = 2v_1$, and $N(A - 5I) = \text{span}\{(1, 2)^\top\}$. For $\lambda_2 = -1$, solve $(A + I)v = 0$:

$$A + I = \begin{pmatrix} 2 & 2 \\ 4 & 4 \end{pmatrix},$$

so both equations reduce to $v_2 = -v_1$, and $N(A + I) = \text{span}\{(1, -1)^\top\}$.

Note that in each case the null space is a line rather than a single vector — any nonzero multiple of an eigenvector is again an eigenvector for the same eigenvalue.

Remark 4.5 (How many eigenvalues). The characteristic polynomial of an $n \times n$ matrix has degree n , so A has at most n distinct real eigenvalues. It can have fewer for two separate reasons: a root may repeat, or a root may fail to be real.

Example 4.6 (A repeated eigenvalue). Let $A = \text{diag}(2, 2)$. Its characteristic polynomial is $(2 - \lambda)^2$, which has the single root 2. Thus A has one distinct eigenvalue rather than two.

Example 4.7 (No real eigenvalues). Let

$$A = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}, \quad \det(A - \lambda I) = \lambda^2 + 1.$$

Since Definition 4.1 requires $\lambda \in \mathbb{R}$, this matrix has no eigenvalues and no nonzero eigenvectors in $\mathbb{R}^2$. It does have complex eigenvalues and eigenvectors, but we do not consider those here.

Below we list some useful results featuring eigenvalues without proving them.

Definition 4.8 (Trace). The *trace* of an $n \times n$ matrix A is the sum of its diagonal entries,

$$\text{tr } A = \sum_{i=1}^n a_{ii}.$$

Theorem 4.9 (Trace and determinant). Let A be an $n \times n$ matrix whose characteristic polynomial factors into real linear terms,

$$\det(A - \lambda I) = (\lambda_1 - \lambda)(\lambda_2 - \lambda) \cdots (\lambda_n - \lambda), \quad \lambda_1, \dots, \lambda_n \in \mathbb{R}.$$

Then

$$\sum_{i=1}^n \lambda_i = \text{tr } A, \quad \prod_{i=1}^n \lambda_i = \det A.$$

Definition 4.10 (Diagonalizable matrix). An $n \times n$ matrix A is *diagonalizable* if there are an invertible $n \times n$ matrix P and a diagonal matrix D with

$$A = PDP^{-1}.$$

Theorem 4.11 (Diagonalization). *An $n \times n$ matrix A is diagonalizable if and only if it has n linearly independent eigenvectors $v^1, \dots, v^n$. In that case one may take*

$$P = [v^1 \ \cdots \ v^n], \quad D = \text{diag}(\lambda_1, \dots, \lambda_n),$$

where λ_i is the eigenvalue belonging to v^i . In particular, A is diagonalizable whenever it has n distinct real eigenvalues.

This result is particularly useful for real symmetric matrices, which always have n linearly independent eigenvectors and are therefore diagonalizable.